

## THE ELLIPTIC CURVE GROUP LAW VIA THE RIEMANN–ROCH THEOREM

### 1. THE PROJECTIVE PLANE

Let  $\mathbb{K}$  be any field. For any nonzero vector  $v = (x, y, z) \in \mathbb{K}^3$ , the *projective equivalence class of  $v$*  is

$$[v] = [x : y : z] = \{\lambda(x, y, z) : \lambda \in \mathbb{K}^\times\}.$$

The  $\mathbb{K}$ -*projective plane* is the set of projective equivalence classes,

$$\mathbb{P}^2\mathbb{K} = \{[x : y : z] : (x, y, z) \in \mathbb{K}^3 - (0, 0, 0)\}.$$

Thus the projective plane is an overlapping union of three sets, each of which can be viewed as a copy of the affine plane translated one unit from the origin in the third direction,

$$\begin{aligned} \mathbb{P}^2\mathbb{K} &= \{[x : y : 1] : (x, y) \in \mathbb{K}^2\} \\ &\cup \{[x : 1 : z] : (x, z) \in \mathbb{K}^2\} \\ &\cup \{[1 : y : z] : (y, z) \in \mathbb{K}^2\}. \end{aligned}$$

Also, the projective plane can be viewed as the disjoint union of three sets, a copy of the affine plane, a line at infinity, and a point at infinity,

$$\begin{aligned} \mathbb{P}^2\mathbb{K} &= \{[x : y : 1] : (x, y) \in \mathbb{K}^2\} \\ &\sqcup \{[x : 1 : 0] : x \in \mathbb{K}\} \\ &\sqcup \{[1 : 0 : 0]\}. \end{aligned}$$

And there are similar disjoint decompositions starting with the  $(x, z)$ -plane or the  $(y, z)$ -plane.

When  $\mathbb{K} = \mathbb{R}$ , our intuition is that the real projective line  $\mathbb{P}^2\mathbb{R}$  is an ordinary line with a point at infinity identifying its opposite directions, and the projective plane is an ordinary plane surrounded by a circle at infinity identifying its opposite directions. The real projective plane differs from the complex projective line —  $\mathbb{P}^1\mathbb{C}$  is an ordinary plane with a single point at infinity gathering all of its directions together.

If  $\mathbb{F}$  and  $\mathbb{K}$  are fields with  $\mathbb{F} \subset \mathbb{K}$  then we can identify  $\mathbb{P}^2\mathbb{F}$  with a subset of  $\mathbb{P}^2\mathbb{K}$ ,

$$\mathbb{P}^2\mathbb{F} \xrightarrow{\sim} \{\mathbb{K}^\times(x, y, z) : [x : y : z] \in \mathbb{P}^2\mathbb{F}\}.$$

That is,  $\mathbb{P}^2\mathbb{F}$  can be viewed as the classes  $[x : y : z]$  in  $\mathbb{P}^2\mathbb{K}$  that have representatives in  $\mathbb{F}^3$ . But not all representatives in  $\mathbb{K}^3$  of such a class lie in  $\mathbb{F}^3$ . For example,  $[i : i : i] = i[1 : 1 : 1] \in \mathbb{P}^2\mathbb{R} \subset \mathbb{P}^2\mathbb{C}$  even though  $i \notin \mathbb{R}$ .

## 2. PLANE CURVES

Now suppose that we have a specific field  $\mathbf{k}$  in mind, and continue to work with three variables. Let  $d$  be a positive integer. A *degree- $d$  curve over  $\mathbf{k}$*  is an irreducible homogeneous polynomial,

$$f(x, y, z) = \sum_{i+j+k=d} a_{ijk} x^i y^j z^k, \quad \text{all } a_{ijk} \in \mathbf{k}.$$

The nonzero solutions in  $\bar{\mathbf{k}}^3$  of the equation  $f(x, y, z) = 0$  form a union of  $\bar{\mathbf{k}}^\times$ -projective equivalence classes, and so we may determine the equation  $f = 0$  as determining a solution set in the projective plane,

$$C_f = \{[x : y : z] \in \mathbb{P}^2 \bar{\mathbf{k}} : f(x, y, z) = 0\}.$$

Note, however, that it is the equation rather than the solution-set that we view as describing the curve. Indeed, for any field  $\mathbb{K}$  such that  $\mathbf{k} \subset \mathbb{K} \subset \bar{\mathbf{k}}$ , the  $\mathbb{K}$ -points of the curve are

$$C_f(\mathbb{K}) = \{[x : y : z] \in \mathbb{P}^2 \mathbb{K} : f(x, y, z) = 0\}.$$

In algebraic geometry the default setting is  $\mathbb{K} = \bar{\mathbf{k}}$ , but in cryptology texts the default setting appears to be  $\mathbb{K} = \mathbf{k}$  instead.

The curve  $f$  is *nonsingular* if its gradient doesn't vanish at any of its points, i.e.,

$$\left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) (x, y, z) \neq (0, 0, 0) \quad \text{for all } (x, y, z) \in C_f.$$

The curve  $f$  has three affine dehomogenizations,  $f(x, y, 1)$ ,  $f(x, 1, z)$ ,  $f(1, y, z)$ . Each dehomogenization determines the intersection of  $C_f$  with a different affine plane. The curve is nonsingular if each of its affine pieces is nonsingular, i.e., if (setting  $z = 1$  and writing  $f(x, y)$  for  $f(x, y, 1)$ )

$$\left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) (x, y) \neq (0, 0) \quad \text{for all } (x, y) \in C_f \cap \{(x, y, 1)\},$$

and similarly for the other two affine pieces of  $f$ .

## 3. ELLIPTIC CURVES

Specialize the discussion to

$$E(x, y, z) = y^2 z - x^3 - axz^2 - bz^3 \quad \text{where } a, b \in \mathbf{k}.$$

(Polynomials of this form are invariably denoted  $E$  rather than  $f$ . From now on we save symbols by letting  $E$  also denote the solution-set, rather than  $C_E$ .) Any nonaffine point  $[x : y : 0]$  of  $E$  satisfies the condition  $x = 0$ , so that in fact the only nonaffine point is  $[0 : 1 : 0]$ . Note that this point lies in  $E(\mathbf{k})$ , i.e., no larger field than  $\mathbf{k}$  is required for the nonaffine point. That is,

$$[0 : 1 : 0] \in E(\mathbb{K}) \quad \text{for any field } \mathbb{K} \text{ between } \mathbf{k} \text{ and } \bar{\mathbf{k}}.$$

Also, the relevant dehomogenization for this point is

$$E(x, 1, z) = z - x^3 - axz^2 - bz^3,$$

with gradient

$$\nabla E(x, 1, z) = (3x^2 - az^2, 1 - 2axz - 3bz^2),$$

so that in particular the gradient at the nonaffine point of  $E$  is nonzero,

$$\nabla E(0, 1, 0) = (0, 1).$$

Thus  $E$  is nonsingular at its nonaffine point.

The affine points of  $E$  are the pairs  $(x, y) \in \bar{\mathbf{k}}^2$  that satisfy the familiar relation defining an elliptic curve,

$$y^2 = x^3 + ax + b.$$

For the affine part of the curve to be nonsingular, the gradient

$$\nabla E(x, y, 1) = (3x^2 + a, 2y)$$

needs to be nonzero at all affine points  $(x, y)$ . Assuming that  $2 \neq 0$  in  $\mathbf{k}$ , the only points of concern are  $(x, 0)$  where  $x^3 + ax + b = 0$ . That is, the curve is nonsingular provided that the cubic polynomial  $c(x) = x^3 + ax + b$  shares no roots in  $\bar{\mathbf{k}}$  with its derivative  $c'(x) = 3x^2 + a$ .

In general, two nonzero polynomials  $f$  and  $g$  over a field  $\mathbf{k}$  share no roots in  $\bar{\mathbf{k}}$  when their *resultant* is nonzero. The resultant is a polynomial function of the coefficients of the two polynomials, the determinant of a matrix called the *Sylvester matrix*. In our case, where  $f(x) = x^3 + ax + b$  and  $g(x) = 3x^2 + a$ , the Sylvester matrix works out to

$$S = \begin{bmatrix} 1 & 0 & a & b & 0 \\ 0 & 1 & 0 & a & b \\ 3 & 0 & a & 0 & 0 \\ 0 & 3 & 0 & a & 0 \\ 0 & 0 & 3 & 0 & a \end{bmatrix},$$

and so the determinant is

$$\begin{vmatrix} 1 & 0 & a & b & 0 \\ 0 & 1 & 0 & a & b \\ 3 & 0 & a & 0 & 0 \\ 0 & 3 & 0 & a & 0 \\ 0 & 0 & 3 & 0 & a \end{vmatrix} = \begin{vmatrix} 1 & 0 & a & b & 0 \\ 0 & 1 & 0 & a & b \\ 0 & 0 & -2a & -3b & 0 \\ 0 & 0 & 0 & -2a & -3b \\ 0 & 0 & 3 & 0 & a \end{vmatrix} = 4a^3 + 27b^2.$$

Thus the condition for nonsingularity is

$$4a^3 + 27b^2 \neq 0 \text{ in } \mathbf{k}.$$

The resultant of a polynomial and its derivative is (up to sign) the *discriminant* of the polynomial. We have just computed the discriminant of a reduced cubic polynomial, where *reduced* means that there is no  $x^2$  term. For a quadratic polynomial, the discriminant works out to the familiar quantity under the square root sign in the quadratic formula, as it must.

#### 4. THE GROUP LAW: EXISTENCE

The elliptic curve

$$E(x, y, z) = y^2z - x^3 - axz^2 - bz^3 \quad \text{where } a, b \in \mathbf{k}, 4a^3 + 27b^2 \neq 0$$

carries a group law, defined entirely in terms of algebraic operations over  $\mathbf{k}$ . Working only in elementary terms, this remarkable fact is difficult to explain satisfactorily.

One partial explanation is that any elliptic curve over the complex number field  $\mathbb{C}$  arises from a quotient of the complex plane by a lattice. This quotient inherits a group structure from the plane, and the group structure carries forward to the curve in  $\mathbb{P}^2\mathbb{C}$ , where it is encapsulated in a slogan:

*Collinear triples add to zero.*

This explanation is augmented by *Bezout's Theorem*, stating that any two curves  $f$  and  $g$  over  $\mathbf{k}$ , of degrees  $d$  and  $e$ , have  $d \cdot e$  points of intersection in  $\mathbb{P}^2\overline{\mathbf{k}}$ . In particular, collinear points on a cubic curve come in triples over any field, not only over  $\mathbb{C}$ . Triples, and only triples, suggest the group law described by the displayed slogan.

Still, the group structure on elliptic curves, and the particular choice of cubic equation used to define elliptic curves, become satisfyingly coherent only once the *Riemann–Roch Theorem* has entered one's life.

The basic principle is that *to understand a mathematical object one needs to understand the functions on the object*. In the case of an elliptic curve — an algebraic object — we need to understand the rational functions, quotients of polynomials in  $x$  and  $y$ . But these behave more subtly than rational functions on the projective line. For example, consider the elliptic curve

$$E(x, y) = y^2 - x^3 + x,$$

and consider the rational function

$$r(x, y) = x/y.$$

Thus  $r(1, 0) = 1/0 = \infty$ , but the nature of  $r(0, 0) = 0/0$  is unclear. However, the relation  $y^2 = x^3 - x$  rewrites as  $x/y = y/(x^2 - 1)$ , so that also  $r(x, y) = y/(x^2 - 1)$  and in particular  $r(0, 0) = 0$ . An rational function on an elliptic curve need not have a single defining formula that works everywhere.

Nonetheless, it is true that the *order of vanishing* of a rational function (excluding the constant function zero) is well defined at each point of  $E$ , including the nonaffine point. The order of vanishing  $\text{ord}_P(r)$  is an integer, as it is for meromorphic functions in complex analysis, and indeed the theory of algebraic curves is partly an outgrowth of the theory of compact Riemann surfaces. Each nonzero rational function on  $E$  has a *divisor*, a finite **formal**  $\mathbb{Z}$ -linear combination of the points of  $E$ ,

$$\text{div}(r) = \sum_{P \in E} \text{ord}_P(r) \cdot (P)$$

In the displayed expression, the symbol  $(P)$  denotes the point  $P$ , the parentheses serving only to remind us that it is appearing in a formal sum rather than an actual sum. That is, no reference is being made to any form of addition on  $E$ , since the idea here is to explain why such addition exists.

A *divisor on  $E$*  is any finite formal  $\mathbb{Z}$ -linear combination of the points of  $E$ ,

$$D = \sum_{P \in E} n_P \cdot (P), \quad \text{all } n_P \in \mathbb{Z}, n_P = 0 \text{ for almost all } P.$$

The set of divisors on  $E$  form an abelian group in the natural way.

The *degree* of a divisor is an integer,

$$\text{deg}(D) = \sum_{P \in E} n_P.$$

The degree map is additive, i.e.,  $\text{deg}(D + \tilde{D}) = \text{deg}(D) + \text{deg}(\tilde{D})$ .

A divisor of the form  $D = \text{div}(r)$  for some rational function  $r$  on  $E$  is called *principal*. All principal divisors have degree 0, but not conversely. The principal divisors form a subgroup of the degree-0 divisors because

$$\text{div}(r) + \text{div}(\tilde{r}) = \text{div}(r\tilde{r}).$$

And so there is a natural quotient group called the *degree-0 Picard group* or the *degree-0 divisor class group*,

$$\text{Pic}^0(E) = \{\text{degree-0 divisors on } E\} / \{\text{principal divisors on } E\}.$$

Every algebraic curve has an associated nonnegative *genus*  $g$ . Linear and quadratic curves have genus  $g = 0$ , elliptic curves, and cubic curves and quadratic curves in general, have genus  $g = 1$ , curves of degree 5 and degree 6 have genus  $g = 2$ , and so on. Thinking of the algebraic curve in complex terms when this is relevant, the genus is also the topological genus, i.e., the complex version of the curve is a  $g$ -holed torus. Nonetheless, the genus is a purely algebraic construct.

Any algebraic curve  $X$  of genus  $g \geq 1$  injects into its degree-0 Picard group. To define such an injection, pick any point  $P_0$  of  $X$  and define a map

$$X \longrightarrow \text{Pic}^0(X), \quad P \longmapsto (P) - (P_0) + \{\text{principal divisors}\}.$$

This is an injection because of the nonobvious fact that no divisor  $(P) - (P_0)$  is principal unless  $P = P_0$ . (In the case of genus  $g = 0$ , every such divisor is principal and the divisor class group is trivial.) And so the question is whether the injection is also a surjection. If it is, then the curve inherits a group structure from its degree-0 Picard group. And the Riemann–Roch Theorem shows, among many other things, that the injection is a surjection if and only if the genus is 1.

To state the Riemann–Roch Theorem, let  $X$  be an algebraic curve. For any divisor  $D = \sum_P n_P \cdot (P)$  on  $X$ , write  $D \geq 0$  if  $n_P \geq 0$  for all  $P$ . Define for any divisor  $D$  on  $X$ ,

$$L(D) = \{0\} \cup \{\text{rational functions } r \text{ on } X \text{ such that } \text{div}(r) + D \geq 0\},$$

and define

$$\ell(D) = \dim(L(D)).$$

Immediately we have:

$$\text{If } \deg(D) < 0 \text{ then } \ell(D) = 0.$$

This is shown by contraposition. If  $\ell(D) > 0$  then there is some nonzero  $r \in L(D)$ . Thus  $\text{div}(r) + D \geq 0$ , and so  $\deg(r) + \deg(D) \geq 0$ . But  $\deg(r) = 0$ , and so  $\deg(D) \geq 0$ .

As one last piece of machinery, we need to invoke that differentials make sense on  $X$ , and they have divisors. A *canonical divisor* is any divisor  $\text{div}(\lambda)$  where  $\lambda$  is a nonzero differential on  $X$ . (Such  $\lambda$  will exist for genus  $g \geq 1$ , not for genus  $g = 0$ .) The Riemann–Roch Theorem says:

*For any divisor  $D$  on  $X$ ,*

$$(1) \quad \boxed{\ell(D) = \deg(D) - g + 1 + \ell(\text{div}(\lambda) - D)}.$$

This is most likely indigestible on first sight. To begin to understand it, we derive some consequences:

- (a)  $\ell(\text{div}(\lambda)) = g$ .
- (b)  $\deg(\text{div}(\lambda)) = 2g - 2$ .
- (c) If  $\deg(D) > 2g - 2$  then  $\ell(D) = \deg(D) - g + 1$ .

To prove (a), set  $D = 0$  in (1) to get  $1 = -g + 1 + \ell(\text{div}(\lambda))$ .

To prove (b), set  $D = \text{div}(\lambda)$  in (1) and quote (a) to get  $g = \deg(\text{div}(\lambda)) - g + 1 + 1$ .

For (c), note that if  $\deg(D) > 2g - 2$  then  $\deg(\operatorname{div}(\lambda) - D) < 0$  by (b), and so  $\ell(\operatorname{div}(\lambda) - D) = 0$  as observed above. And we are done by the Riemann–Roch formula (1).

With the consequences in place we can rewrite the Riemann–Roch formula (1) in a symmetric form that no longer makes direct reference to the genus but shows that the Riemann–Roch Theorem is a duality theorem,

$$\ell(D) - \frac{1}{2} \deg(D) = \ell(D') - \frac{1}{2} \deg(D'), \quad \text{where } D' = \operatorname{div}(\lambda) - D.$$

For an application of the Riemann–Roch Theorem, consider a curve  $X$  of genus  $g = 1$ . Let  $P$  be a point of  $X$ . By (c),

$$\ell(n \cdot (P)) = n \quad \text{for } n = 1, 2, 3, \dots$$

Thus  $\ell((P)) = 1$  and so  $L((P)) = \bar{\mathbf{k}}$ , i.e.,

$$L((P)) \text{ has basis } \{1\}.$$

And  $\ell(2(P)) = 2$  and so  $L(2(P)) = \bar{\mathbf{k}} \oplus \bar{\mathbf{k}}x$  where  $\operatorname{ord}_P(x) = -2$  and  $\operatorname{ord}_Q(x) \geq 0$  for all other points  $Q$ ,

$$L(2(P)) \text{ has basis } \{1, x\}.$$

And  $\ell(3(P)) = 3$  and so  $L(3(P)) = \bar{\mathbf{k}} \oplus \bar{\mathbf{k}}x \oplus \bar{\mathbf{k}}y$  where  $\operatorname{ord}_P(y) = -3$  and  $\operatorname{ord}_Q(y) \geq 0$  for all other points  $Q$ ,

$$L(3(P)) \text{ has basis } \{1, x, y\}.$$

Similarly,

$$L(4(P)) \text{ has basis } \{1, x, y, x^2\},$$

and

$$L(5(P)) \text{ has basis } \{1, x, y, x^2, xy\}.$$

But things get interesting at  $6(P)$ : both  $x^3$  and  $y^2$  have order  $-6$  at  $P$  and non-negative order everywhere else. Therefore:

The set  $\{1, x, y, x^2, xy, x^3, y^2\}$  is linearly dependent,

and the linear relation must involve both  $x^3$  and  $y^2$ . That is, for appropriate constants, after suitable normalization,

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Assuming that  $2 \neq 0$  and  $3 \neq 0$ , further substitutions reduce the equation to the familiar formula  $y^2 = x^3 + ax + b$ . The Weierstrass relation has emerged naturally from the Riemann–Roch Theorem.

Another argument with the Riemann–Roch Theorem and further machinery shows that for genus  $g = 1$ , the embedding of a curve into its degree-0 Picard group is indeed a surjection, and hence the curve carries a group structure.

One reason that it took people so long to notice the Picard group (or an isomorphic construction, the *Jacobian* of a curve), is that the object is trivial for genus  $g = 0$  and naturally isomorphic to the curve itself for genus  $g = 1$ . Only for higher genus  $g \geq 2$  do the Picard group and the Jacobian emerge clearly as objects in their own right.

## 5. THE GROUP LAW: FORMULAS

To state the formulas for the group law, recall that for any field  $\mathbb{K}$  between  $\mathbf{k}$  and  $\bar{\mathbf{k}}$  we have

$$E(\mathbb{K}) = \{(x, y) \in \mathbb{K}^2 : y^2 = x^3 + ax + b\} \cup \{\mathcal{O}\} \quad \text{where } \mathcal{O} = [0 : 1 : 0].$$

Consider any two points  $p_1, p_2 \in E(\mathbb{K})$ . The group law is

- If  $p_1 = \mathcal{O}$  then  $p_1 + p_2 = p_2$ , and if  $p_2 = \mathcal{O}$  then  $p_1 + p_2 = p_1$ .

For the rest of the cases, take  $p_1 = (x_1, y_1) \neq \mathcal{O}$  and  $p_2 = (x_2, y_2) \neq \mathcal{O}$ . Note that if  $x_1 = x_2$  then either  $y_1 = -y_2$  (this includes the case  $y_1 = y_2 = 0$ ) or  $y_1 = y_2 \neq 0$ .

- If  $x_1 = x_2$  and  $y_1 = -y_2$  then  $p_1 + p_2 = \mathcal{O}$ .
- Otherwise set

$$m = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & \text{if } x_1 \neq x_2, \\ \frac{3x_1^2 + a}{2y_1} & \text{if } x_1 = x_2 \end{cases}$$

and

$$x_3 = m^2 - x_1 - x_2.$$

Then

$$p_1 + p_2 = (x_3, -m(x_3 - x_1) - y_1).$$

Consequently the additive inverse of any affine point  $(x, y)$  of  $E(\mathbb{K})$  is  $(x, -y)$ .

## 6. AN ELLIPTIC CURVE VERSION OF ELGAMAL

The ElGamal cipher for elliptic curves is perfectly analogous to the ElGamal cipher for  $(\mathbb{Z}/p\mathbb{Z})^\times$ . The only difference is notational: since the group law on an elliptic curve is written additively while the group law on  $(\mathbb{Z}/p\mathbb{Z})^\times$  is written multiplicatively, the elliptic curve ElGamal description appears to be the “logarithm” of the  $(\mathbb{Z}/p\mathbb{Z})^\times$  ElGamal description.

Alice chooses an elliptic curve  $E$  over the field  $\mathbf{k} = \mathbb{Z}/p\mathbb{Z}$  where  $p$  is a large prime. She chooses a point  $a$  on  $E$  and a positive integer key  $\ell$ . She computes  $b = \ell \cdot a$  (meaning  $a$  added to itself  $\ell$  times). She publishes  $E$ ,  $p$ ,  $a$ , and  $b$ , but she keeps  $\ell$  secret.

Bob can then send a message that only Alice can decrypt. His plaintext is a point  $\pi \in E(\mathbf{k})$ . He chooses an auxiliary random integer  $r$ , and he encrypts a header

$$r \cdot a \in E(\mathbf{k}),$$

and he encrypts the plaintext  $\pi$  as the ciphertext

$$\eta = \pi + r \cdot b \in E(\mathbf{k}).$$

His ciphertext  $\eta$  masks the plaintext  $\pi$ , while his header  $r \cdot a$  gives Alice the necessary information to unmask it.

To decrypt, Alice computes from the header

$$\ell \cdot (r \cdot a) = r \cdot (\ell \cdot a) = r \cdot b,$$

and then she computes the additive inverse  $-(r \cdot b)$ . After that, she computes from the plaintext,

$$-(r \cdot b) + \eta = -(r \cdot b) + \pi + r \cdot b = \pi.$$

## 7. AN ELLIPTIC CURVE VERSION OF DIFFIE-HELLMAN

Alice and Bob (and everyone else) know a prime  $p$ , and an elliptic curve  $E$  over the field  $\mathbf{k} = \mathbb{Z}/p\mathbb{Z}$ , and a point  $\pi$  of  $E(\mathbf{k})$ . Alice randomly chooses a secret positive integer  $n_A$ , and similarly Bob randomly chooses a secret positive integer  $n_B$ . Alice sends the point  $\pi_A = n_A \cdot \pi \in E(\mathbf{k})$  to Bob over an insecure channel, while Bob sends  $\pi_B = n_B \cdot \pi$  to Alice. Alice computes

$$n_A \cdot \pi_B = n_A \cdot (n_B \cdot \pi) = n_A n_B \cdot \pi,$$

and Bob computes

$$n_B \cdot \pi_A = n_B \cdot (n_A \cdot \pi) = n_A n_B \cdot \pi.$$

This is their shared key.

## 8. ELLIPTIC CURVE FACTORING

Suppose that we have an integer  $n$  that is known to be of the form  $n = pq$ , but we don't know  $p$  and  $q$ . Choose a random cubic polynomial over  $\mathbb{Z}/n\mathbb{Z}$ ,

$$E(x, y) = y^2 - x^3 - ax - b, \quad a, b \in \mathbb{Z}/n\mathbb{Z}, \quad 4a^3 + 27b^2 \not\equiv 0 \pmod{n},$$

and choose a point  $\pi$  on  $E(\mathbb{Z}/n\mathbb{Z})$ . (In practice, the easy way to do this is to choose the first polynomial coefficient  $a$  and the point  $\pi = (x, y)$ , and then to choose the remaining polynomial coefficient  $b$  accordingly.)

By the Sun-Ze Theorem, tracking data modulo  $n$  is the same as tracking compatible pairs of data modulo  $p$  and modulo  $q$ , even though we don't know  $p$  and  $q$ . So the idea is to use the group law formulas to compute

$$2 \cdot \pi, \quad 3 \cdot \pi, \quad 4 \cdot \pi, \quad \dots$$

until the formula

$$m = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & \text{if } x_1 \neq x_2, \\ \frac{3x_1^2 + a}{2y_1} & \text{if } x_1 = x_2 \end{cases}$$

fails because  $x_2 - x_1$  or  $2y_1$  fails to be invertible modulo  $n$ . This will happen exactly when

$$\gcd(x_2 - x_1, n) > 1 \quad \text{or} \quad \gcd(2y_1, n) > 1.$$

That is, one but not both of  $p, q$  divides the denominator. The idea is that running the group law modulo  $n$  is effectively computing compatible point-multiples on two elliptic curves, one defined modulo  $p$ , the other modulo  $q$ . The original point  $\pi$  defines points  $\pi_p \in E_1(\mathbb{Z}/p\mathbb{Z})$  and  $\pi_q \in E_2(\mathbb{Z}/q\mathbb{Z})$ . There is no reason to believe that  $\pi_p$  and  $\pi_q$  should have the same order in their respective abelian groups, and the process is computing until some multiple of (say)  $\pi_p$  is zero, but not the corresponding multiple of  $\pi_q$ . This exposes  $p$ .

This factoring method is the elliptic curve analogue of Pollard's  $p - 1$  method. Just as the  $p - 1$  method did not work well when the  $p - 1$  has a large prime factor, the elliptic curve method does not obviously work well when the order of  $\pi_p$  and of  $\pi_q$  both have large prime factors. However, whereas there is only *one*  $(\mathbb{Z}/n\mathbb{Z})^\times$ , there are *many* elliptic curves over  $\mathbb{Z}/n\mathbb{Z}$ . One can run the algorithm using many curves in parallel, or one can start over with a new curve if the chosen curve doesn't yield results after a while.